

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-184085

(43)Date of publication of application : 06.07.2001

(51)Int.Cl. G10L 15/22
G10L 15/00
H04Q 7/38

(21)Application number : 11-364091

(71)Applicant : NEC SAITAMA LTD

(22)Date of filing : 22.12.1999

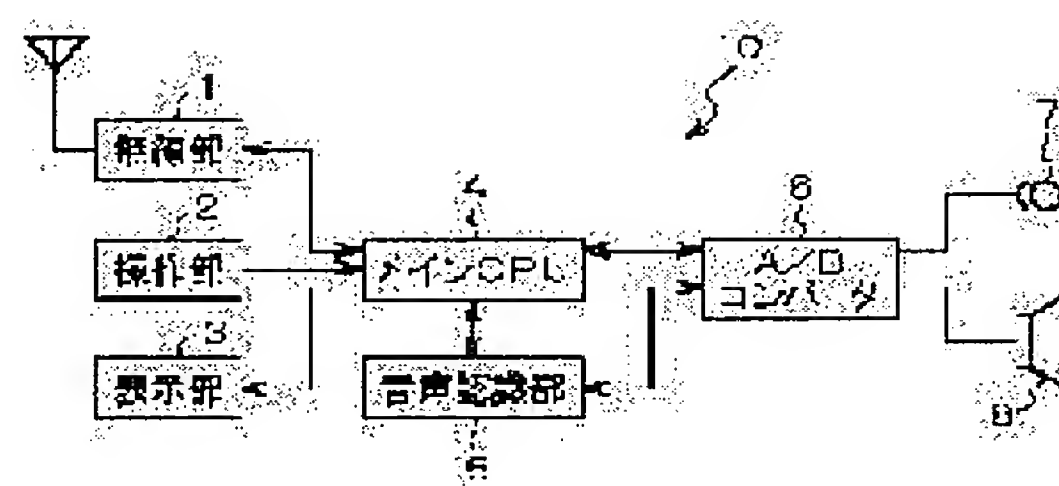
(72)Inventor : SHIONO KATSUMI

(54) VOICE RECOGNITION DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To reduce misrecognition due to the utterance of an unregistered word and noise.

SOLUTION: Voice recognition is carried out by generating a misrecognition countermeasure word based on a registered word and deciding the word having the higher similarity to the registered word between the registered word and misrecognition counter measure word as a recognition result; when the misrecognition word is the misrecognition countermeasure word, misrecognition is decided. The registered word is attenuated with the misrecognition countermeasure word and the utterance of a word which is not a registered word and the utterance of a registered word in noisy environment bring out misrecognition with high possibility. The vowel of the syllabary of an n-th character of the misrecognition countermeasure word is preferably a vowel with low appearance frequency in the syllabary of the n-th character of the registered word. A misrecognition countermeasure word like this has low similarity to the registered word and when a user voices the registered word, the possibility that the misrecognition countermeasure word is a recognition result is low.



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開2001－184085

(P2001－184085A)

(43)公開日 平成13年 7 月 6 日 (2001. 7. 6)

| (51)Int.Cl. ⁷ | 識別記号 | F I | テーマコード* (参考) | |
|----------------------------|------|--------------|--------------|-----------|
| G 1 0 L 15/22 | | G 1 0 L 3/00 | 5 6 1 F | 5 D 0 1 5 |
| 15/00 | | | 5 5 1 A | 5 K 0 6 7 |
| H 0 4 Q 7/38 | | | 5 6 1 D | 9 A 0 0 1 |
| | | H 0 4 B 7/26 | 1 0 9 H | |
| 審査請求 有 請求項の数 8 O L (全 5 頁) | | | | |

(21)出願番号 特願平11－364091

(22)出願日 平成11年12月22日 (1999. 12. 22)

(71)出願人 390010179

埼玉日本電気株式会社

埼玉県児玉郡神川町大字元原字豊原300番
18

(72)発明者 塩野 勝美

埼玉県児玉郡神川町大字元原字豊原300番
18 埼玉日本電気株式会社内

(74)代理人 100102864

弁理士 工藤 実 (外 1 名)

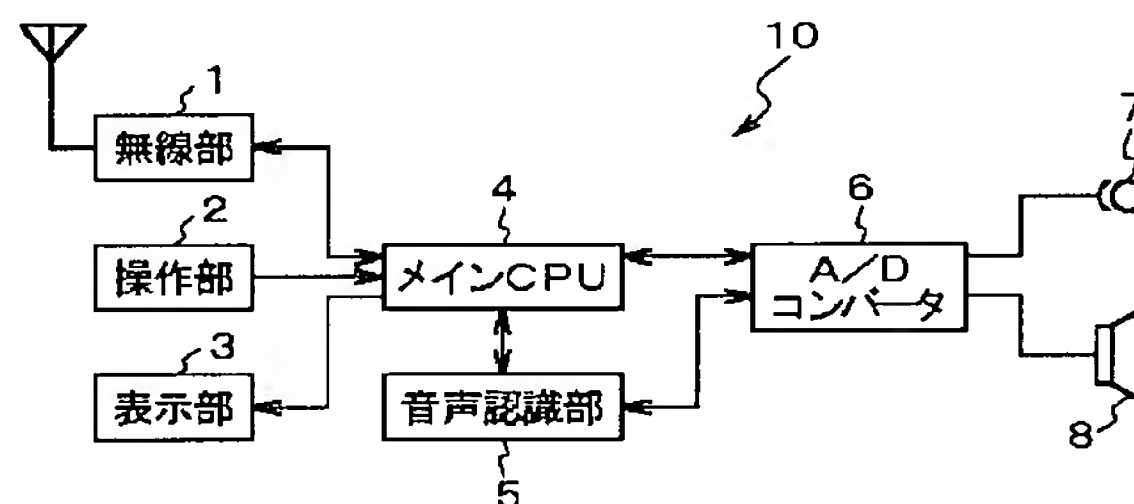
Fターム(参考) 5D015 KK01 KK04 LL04 LL05 LL06
5K067 AA33 BB04 EE02 HH22 HH23
9A001 EZ05 HH17

(54)【発明の名称】 音声認識装置

(57)【要約】

【課題】登録されていない単語の発声、騒音による誤認識をより低減させる。

【解決手段】登録された登録単語に基づいて誤認識対策単語を作成し、登録単語と誤認識対策単語との中で、発声され入力された発声単語に類似度が最も高い単語を認識結果とする音声認識を行い、認識結果が、誤認識対策単語である場合、誤認識とする。登録単語を誤認識対策単語で希釈することにより、登録単語でない単語の発声、騒音の中での登録単語の発声は、誤認識となる確率が高くなる。誤認識対策単語のn文字目の音節の母音は、登録単語のn文字目の音節において出現頻度の低い母音であることが好ましい。このような誤認識対策単語は、登録単語との類似度が低く、使用者が登録単語を発声した場合、誤認識対策単語が認識結果となる確率は低くなる。



【特許請求の範囲】

【請求項 1】登録された登録単語に基づいて誤認識対策単語を作成する作成手段と、
前記登録単語と前記誤認識対策単語との中で、入力された発声単語に類似度が最も高い単語を認識結果とする音声認識手段とを含み、
前記認識結果が、前記誤認識対策単語である場合、誤認識とする音声認識装置。

【請求項 2】請求項 1 において、
前記誤認識対策単語の n 文字目の音節の母音は、前記登録単語の n 文字目の音節では出現頻度の低い母音であり、
音声認識装置。

【請求項 3】請求項 2 において、
前記誤認識対策単語の n 文字目の音節の子音は、前記登録単語の n 文字目の音節では出現頻度の低い子音である
音声認識装置。

【請求項 4】請求項 2 において、更に、
スピーカを含み、
前記スピーカは、前記誤認識を使用者に通知する音声認識装置。

【請求項 5】請求項 2 において、更に、
表示部を含み、
前記表示部は、前記誤認識を使用者に通知する音声認識装置。

【請求項 6】請求項 2 において、
前記誤認識は、再度、前記音声認識手段を起動させる音声認識装置。

【請求項 7】請求項 2 において、
前記誤認識対策単語は、前記登録単語が、新たに登録される場合に作成され、
前記誤認識対策単語の音節数は、前記登録単語の中の 1 単語の音節数と等しく、
前記誤認識対策単語は、音節数毎に複数を作成される音声認識装置。

【請求項 8】請求項 2 において、
前記音声認識装置は、携帯電話機である音声認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声認識装置に関し、特に、登録されていない単語の発声、騒音による誤認識をより低減させる音声認識装置に関する。

【0002】

【従来の技術】マイクから入力された音声を認識し、認識した単語に対応した処理を行う携帯電話機が提案されている。音声認識処理は、単語音声を母音と子音の組み合わせからなる最小単位の音節に分け、各音節単位で認識を行う処理が理想である。しかし、実際に発声された単語音声を音節単位に区切ることは、非常に困難である。そこで、認識させる単語を予め登録しておき、発声

された音声と全ての登録単語との間で比較処理を行い、登録単語の中で最も類似度の高い単語を認識結果とする方法が採用されている。この場合、使用者が誤って登録されていない単語を発声したり、周囲に騒音がある所で発声したりすると、意図しない登録単語に誤認識されてしまうことがある。

【0003】

【発明が解決しようとする課題】本発明の課題は、登録されていない単語の発声、騒音による誤認識をより低減させる音声認識装置を提供することにある。

【0004】

【課題を解決するための手段】その課題を解決するための手段が、下記のように表現される。その表現中に現れる技術的事項には、括弧（）付きで、番号、記号等が添記されている。その番号、記号等は、本発明の実施の複数・形態又は複数の実施例のうちの少なくとも 1 つの実施の形態又は複数の実施例を構成する技術的事項、特に、その実施の形態又は実施例に対応する図面に表現されている技術的事項に付せられている参照番号、参照記号等に一致している。このような参照番号、参照記号は、請求項記載の技術的事項と実施の形態又は実施例の技術的事項との対応・橋渡しを明確にしている。このような対応・橋渡しは、請求項記載の技術的事項が実施の形態又は実施例の技術的事項に限定されて解釈することを意味しない。

【0005】本発明による音声認識装置は、登録された登録単語に基づいて誤認識対策単語を作成する作成手段と、登録単語と誤認識対策単語との中で、入力された発声単語に類似度が最も高い単語を認識結果とする音声認識手段とを含み、認識結果が、誤認識対策単語である場合、誤認識とする。登録単語を誤認識対策単語で希釈することにより、登録単語でない単語の発声、騒音の中での登録単語の発声は、誤認識対策単語が認識結果となる確率が高くなる。

【0006】誤認識対策単語の n 文字目の音節の母音は、登録単語の n 文字目の音節において出現頻度の低い母音であることが好ましい。このような誤認識対策単語は、類似度が低く、使用者が登録単語を発声した場合、誤認識対策単語が認識結果となる確率は低くなる。誤認識対策単語の n 文字目の音節の子音は、登録単語の n 文字目の音節において出現頻度の低い子音である。このような誤認識対策単語は、母音のみによって作成された単語よりも、より類似度が低くなる。

【0007】更に、スピーカを含み、スピーカは、誤認識を使用者に通知する。更に、表示部を含み、表示部は、誤認識を使用者に通知する。または、誤認識は、音声認識を再起動させる。このような機能は、使い勝手の向上の点で好ましい。

【0008】誤認識対策単語は、登録単語が、新たに登録される場合に作成され、誤認識対策単語の音節数は、

登録単語の中の1つの音節数と等しいことが好ましい。

【0009】音声認識装置は、携帯電話機である。登録単語と携帯電話機の機能とを対応づけることにより、携帯電話の使い勝手が良くなる。

【0010】

【発明の実施の形態】図に一致対応して、本発明による音声認識装置の実施の形態は、携帯電話機である。図1に示されるように、携帯電話機10は、無線部1、操作部2、表示部3、メインCPU4、音声認識部5、A/Dコンバータ6、マイク7及びスピーカ8からなる。無線部1には、アンテナが付属している。無線部1は、メインCPU4と接続している。操作部2は、メインCPU4と接続している。操作部2は、図示されていない音声認識開始キーを有している。メインCPU4は、表示部3、音声認識部5及び、A/Dコンバータ6と接続している。音声認識部5は、A/Dコンバータ6を接続している。マイク7は、A/Dコンバータ6と接続している。A/Dコンバータ6は、スピーカ8と接続している。

【0011】無線部1は、図示されていない基地局との無線信号の送受信を行う。操作部2は、使用者により携帯電話機の操作、登録する単語の入力、音声認識の開始の指示が行われる。表示部3は、メインCPU4の指示により、数字、文字などを表示する。マイク7は、音声が入力される。A/Dコンバータ6は、音声データのアナログデジタル変換を行う。音声認識部5は、デジタル変換された音声データにより、音声認識処理を行う。スピーカ8は、音声認識時の開始音と、認識結果に対応する音声とを出力する。メインCPU4は、携帯電話全体の制御を行う。

【0012】図2は、使用者が新たな単語を登録するときの処理フローを示している。フローは、ステップ21～24からなる。ステップ21では、使用者により単語の登録が行われたかどうか判定される。使用者により単語と、その単語に対応する処理とが操作部2に入力される。操作部2は、単語と、その単語に対応する処理とをメインCPU4に出力する。メインCPU4は、操作部2から出力された単語と処理とを登録する。登録が行われた場合、ステップ22に移行する。登録が行われなかった場合、再度フローはステップ21に戻る。

【0013】ステップ22では、既に登録済みの単語に新たに登録された単語を加えて、登録済み単語の統計処理を行う。例えば、登録済みの単語として、図3の表に示されるように、イシイ、カトウ、ナカムラ、オノ、ヤマグチの5単語があるとする。各単語のn文字目の音節の母音について整理する。登録単語1のイシイの場合、母音は、1文字目がi、2文字目がi、3文字目がiとなる。図4の表に示されるように、全ての単語のn文字目の母音の累計を計算する。本例では、1文字目の母音の累計は、aが3個、iが1個、u、eが0個、oが1

個のようになる。2文字目から4文字目までについても同様に母音の累計を計算する。

【0014】ステップ23では、登録された全単語のn文字目の母音の累計結果を用いて、誤認識対策単語を作成する。誤認識対策単語は、登録済みの単語と同じ長さの単語を、2種類ずつ作成される。類似度が低い単語となる母音は、累計が少ない母音である。1文字目の母音の累計が0個であるu、eが作成する単語の1文字目の母音となる。同様にして、2文字目の母音は、u、e、3文字目の母音は、a、e、o、4文字目の母音は、u、e、oとなる。2文字の単語を作成する場合、母音は1文字目が、u、e、2文字目がu、eとなり、例えば、「ウエ」、「エウ」が作成される。登録済みの単語は2、3、4文字の単語が存在している。3、4文字の単語についても、2文字の単語と同様に作成される。

【0015】ステップ24では、メインCPU4により作成された単語を登録する。本例では、作成された単語の長さを登録済みの単語と同じ長さとしたが、誤認回避のため、単語の長さを変更してもよい。作成する単語の個数は、2個としたが、個数を変更してもよい。

【0016】図5は、本発明による音声認識装置の音声認識のフローを示している。フローはステップ31～35からなる。操作部2が有する音声認識開始キーが押下されると、フローは起動する。ステップ31では、メインCPU4が、音声認識開始キーの押下を検出し、A/Dコンバータ6を介して、スピーカ8に、音声認識の開始音の鳴動を指示する。スピーカ8は、開始音を出力する。その後、メインCPU4は、音声認識の開始を音声認識部5に命令する。

【0017】ステップ32では、入力された音声の認識処理が行われる。音声認識部5は、音声認識の開始を命令されると、音声データ入力待ちの状態になる。マイク7は、入力された音声を、A/Dコンバータ6に出力する。A/Dコンバータ6は、マイク7の出力をデジタルデータである音声データに変換し、音声認識部5に音声データを出力する。音声データを受け取った音声認識部5は、認識結果を登録済みの単語と比較する。認識結果が登録済みの単語の中にあれば、音声認識部5は、認識結果をメインCPU4に出力し、フローは、ステップ33に移行する。認識結果が登録済みの単語の中になければ、フローは、再度、ステップ32に移行し、音声認識部5は、音声データ入力待ちの状態になる。

【0018】ステップ33では、認識結果がどのように登録された単語かを判定する。認識結果を受け取ったメインCPU4は、認識結果がメインCPU4により登録された単語か、使用者により登録された単語かを判別する。認識結果が、メインCPU4により登録された単語である場合、ステップ34に移行する。認識結果が使用者により登録された単語である場合、ステップ35に移行する。

【0019】ステップ34では、音声を誤認識したことを使用者に通知する。認識結果がメインCPU4により作成された単語であることは、使用者が誤って登録していない単語を発声し、または、周囲の騒音により誤認識したと解釈される。メインCPU4は、誤認識した旨の表示を表示部3に指示する。表示部3は、メインCPU4の指示により、誤認識した旨を通知するメッセージを表示する。メインCPU4は、誤認識した旨の音声の出力をスピーカ8に指示する。スピーカ8は、メインCPU4の指示により、誤認識した旨の音声出力する。

【0020】ステップ35では、使用者により登録された単語に対応した処理が行われる。メインCPU4は、使用者により登録された単語に対応するメッセージを表示部3に表示させ、対応する音声をスピーカ8に出力させる。

【0021】従来の音声認識装置では、騒音がある所で発声したり、使用者が誤って未登録の単語を発声したりすると、意図しない登録単語に誤認識されやすい。本発明による音声認識装置によれば、騒音がある所で発声された登録単語や、未登録の単語は、誤認識対策単語と認識されて誤認識を回避できる。

【0022】使用者自らが誤動作を防止するために単語を登録しておくことで、誤認識の低減を図ることもできる。登録単語数が多くなると、使用者が誤って登録した単語との類似性が高い単語を作成してしまうことがある。この場合、使用者が登録済みの単語を発声したにもかかわらず、誤動作対策用の単語に認識される恐れがある。本発明による音声認識装置によれば、誤動作を防止するための誤認識対策単語を使用者自らが作成する必要がない。

【0023】本発明による音声認識装置の実施の他の形態は、先の実施の形態における音声認識処理方法が改められている。

【0024】図6は、本発明による音声認識装置の音声認識のフローを示している。フローはステップ41～44からなる。操作部2が有する音声認識開始キーが押下されると、フローは起動する。ステップ41では、メインCPU4が、音声認識開始キーの押下を検出し、A/Dコンバータ6を介して、スピーカ8に、音声認識の開始音の鳴動を指示する。スピーカ8は、開始音を出力する。その後、メインCPU4は、音声認識の開始を音声認識部5に命令する。

【0025】ステップ42では、入力された音声の認識処理が行われる。音声認識部5は、音声認識の開始を指示されると、音声データ入力待ちの状態になる。マイク7は、入力された音声を、A/Dコンバータ6に出力する。A/Dコンバータ6は、マイク7の出力をデジタルデータである音声データに変換し、音声認識部5に音声データを出力する。音声データを受け取った音声認識部5は、認識結果を登録済みの単語と比較する。認識結果

が登録済みの単語の中になれば、音声認識部5は、認識結果をメインCPU4に出力し、フローは、ステップ43に移行する。認識結果が登録済みの単語の中になければ、フローは、再度、ステップ42に移行し、音声認識部5は、音声データ入力待ちの状態になる。

【0026】ステップ43では、認識結果がどのように登録された単語かを判定する。認識結果を受け取ったメインCPU4は、認識結果がメインCPU4により作成された単語か、使用者により登録された単語かを判別する。認識結果が、メインCPU4により登録された単語である場合、フローは、再度、ステップ42に移行し、音声認識部5は、音声データ入力待ちの状態になる。認識結果が使用者により登録された単語である場合、ステップ44に移行する。

【0027】ステップ44では、使用者により登録された単語に対応した処理が行われる。メインCPU4は、使用者により登録された単語に対応するメッセージを表示部3に表示させ、対応する音声をスピーカ8に出力させる。

【0028】本発明による音声認識装置の本実施の形態によれば、誤動作が発生した場合、再度、使用者は音声認識を開始させる必要がなく使い勝手が良い。

【0029】本発明による音声認識装置の実施の更に他の形態は、先の実施の形態における類似度が低い単語の作成方法が改められている。メインCPU4は、登録済みの単語の母音情報に加えて、子音情報の統計処理を行い、類似度が低い単語を作成する。こうすることにより、類似度がより低い単語を作成することができる。

【0030】

【発明の効果】本発明による音声認識装置は、登録されていない単語を使用者が誤って発声したり、音声認識中に周囲の騒音があつたりする場合の誤認識を、より低減させることができる。

【図面の簡単な説明】

【図1】図1は、本発明による音声認識装置の実施の形態を示すブロック図である。

【図2】図2は、本発明による音声入力方法の実施の形態における誤認識対策単語の作成方法を示すフローチャートである。

【図3】図3は、本発明による音声入力方法の実施の形態における認識辞書の構成例を示す表である。

【図4】図4は、本発明による音声入力方法の実施の形態における登録単語のn文字目の母音の累計を示す表である。

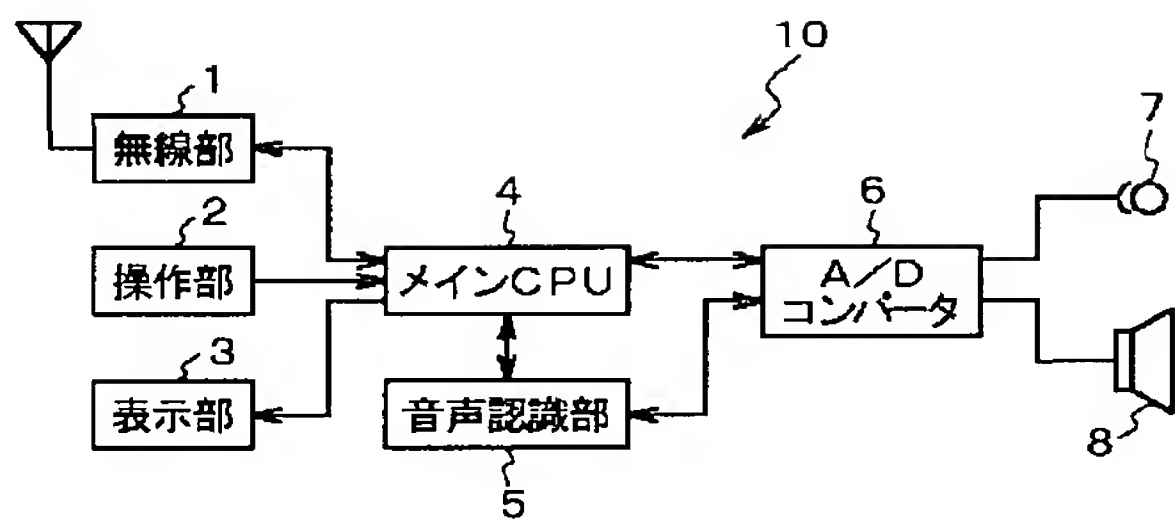
【図5】図5は、本発明による音声入力方法の実施の形態における音声認識処理を示すフローチャートである。

【図6】図6は、本発明による音声入力方法の実施の他の形態における音声認識処理を示すフローチャートである。

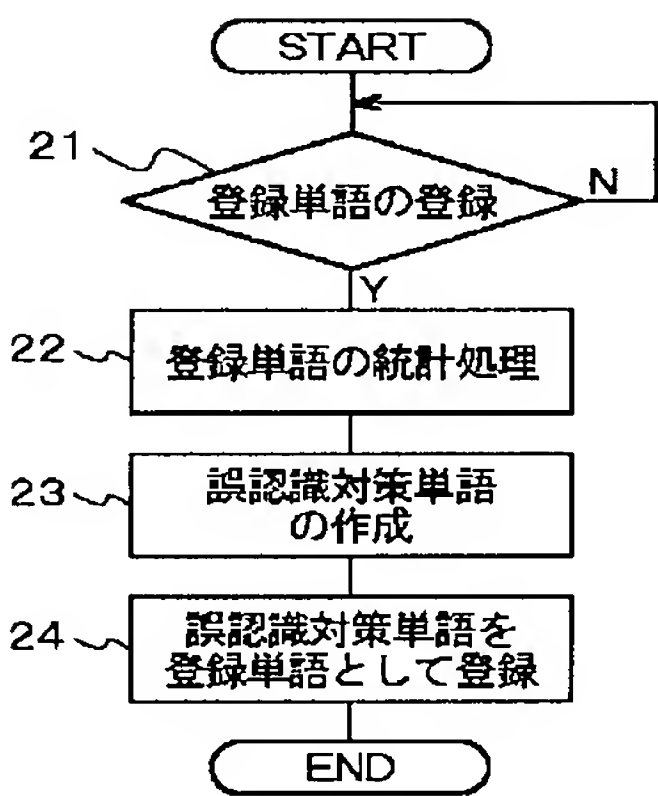
【符号の説明】

- 1…無線部
2…操作部
3…表示部
4…メインCPU
5…音声認識部
- * 6…A/Dコンバータ
7…マイク
8…スピーカ
10…携帯電話機
- *

【図1】



【図2】



【図3】

＜ 認識辞書の構成例 ＞

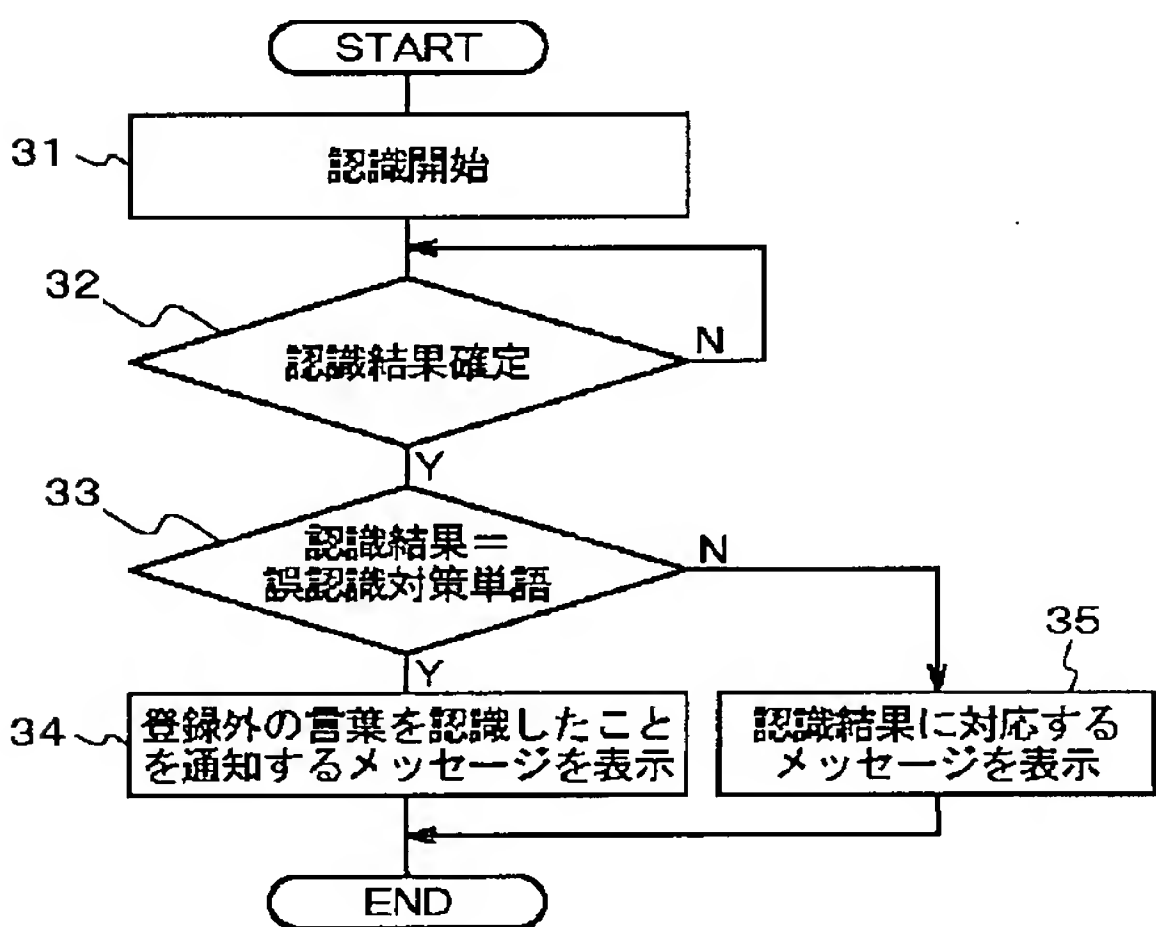
| | 登録単語 | 登録単語の母音 |
|-------|------|---------|
| 登録単語1 | イシイ | i i i |
| 登録単語2 | カトウ | a o u |
| 登録単語3 | ナカムラ | a a u a |
| 登録単語4 | オノ | o o |
| 登録単語5 | ヤマグチ | a a u i |

【図4】

＜ n文字目の母音の累計 ＞

| 母音の種類 | 1文字目 | 2文字目 | 3文字目 | 4文字目 |
|-------|------|------|------|------|
| a | 3 | 2 | 0 | 1 |
| i | 1 | 1 | 1 | 1 |
| u | 0 | 0 | 3 | 0 |
| e | 0 | 0 | 0 | 0 |
| o | 1 | 2 | 0 | 0 |

【図5】



【図6】

